

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/646,390A

Source: FFWO

Date Processed by STIC: 11-15-04

**ENTERED**



IFWO

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/646,390A

DATE: 11/15/2004

TIME: 12:00:48

Input Set : A:\PTO.FG.txt

Output Set: N:\CRF4\11152004\J646390A.raw

( pg. 6 )

3 <110> APPLICANT: SAITOU, Mitinori  
 4 SURANI, Azim  
 6 <120> TITLE OF INVENTION: Genes  
 8 <130> FILE REFERENCE: 674558-2002.1  
 10 <140> CURRENT APPLICATION NUMBER: 10/646,390A  
 11 <141> CURRENT FILING DATE: 2003-08-21  
 13 <150> PRIOR APPLICATION NUMBER: 10/621,911  
 14 <151> PRIOR FILING DATE: 2003-07-17  
 16 <150> PRIOR APPLICATION NUMBER: PCT/GB02/00215  
 17 <151> PRIOR FILING DATE: 2002-01-18  
 19 <150> PRIOR APPLICATION NUMBER: GB 0101300.2  
 20 <151> PRIOR FILING DATE: 2001-01-18  
 22 <160> NUMBER OF SEQ ID NOS: 26  
 24 <170> SOFTWARE: SeqWin99, version 1.02  
 26 <210> SEQ ID NO: 1  
 27 <211> LENGTH: 617  
 28 <212> TYPE: DNA  
 29 <213> ORGANISM: Mus musculus  
 31 <400> SEQUENCE: 1

32 gccgcagaaa gggcagacccc gcagcgcgct ccatcctttg ccctccagtg ctgcctttgc 60  
 33 tccgcaccat gaaccacact tctcaaggct tcataccgc tgccagtgga ggacagcccc 120  
 34 caaaactacga aagaatcaag gaagaatatg aggtggctga gatgggggca ccgcacggat 180  
 35 cggcttctgt cagaactact gtgatcaaca tgcccagaga ggtgtcggtg cctgaccatg 240  
 36 tggctctggtc cctgttcaat acacttctca tgaacttctg ctgcctgggc ttcatagcct 300  
 37 atgcctactc cgtaagact agggatcggg agatgggtggg tgatgtgact ggagcccccagg 360  
 38 cctacgcctc cactgctaag tgcctgaaca tcagcacctt ggtcctcagc atccctgatgg 420  
 39 ttgttatcac cattgttagt gtcatacatca ttgttcttaa cgctcaaaac cttcacactt 480  
 40 aatagaggat tccgacttcc ggtcctgaag tgcttcaccc tccgcagctg cgtccctcct 540  
 41 tgcccctccc tacacgcagg tgtaacactc atttatctat ccacagtggta ttcaataaaag 600  
 42 tgcacttgat aaccacc 617

44 <210> SEQ ID NO: 2  
 45 <211> LENGTH: 137  
 46 <212> TYPE: PRT  
 47 <213> ORGANISM: Mus musculus  
 49 <400> SEQUENCE: 2

50	Met	Asn	His	Thr	Ser	Gln	Ala	Phe	Ile	Thr	Ala	Ala	Ser	Gly	Gly	Gln
51	1															
52																
53																
54																
55																
56																
57																
58																
59																
60																

50 Met Asn His Thr Ser Gln Ala Phe Ile Thr Ala Ala Ser Gly Gly Gln  
 51 1 5 10 15  
 52  
 53 Pro Pro Asn Tyr Glu Arg Ile Lys Glu Glu Tyr Glu Val Ala Glu Met  
 54 20 25 30  
 55  
 56 Gly Ala Pro His Gly Ser Ala Ser Val Arg Thr Thr Val Ile Asn Met  
 57 35 40 45  
 58  
 59 Pro Arg Glu Val Ser Val Pro Asp His Val Val Trp Ser Leu Phe Asn  
 60 50 55 60

RAW SEQUENCE LISTING DATE: 11/15/2004  
 PATENT APPLICATION: US/10/646,390A TIME: 12:00:48

Input Set : A:\PTO.FG.txt  
 Output Set: N:\CRF4\11152004\J646390A.raw

62 Thr Leu Phe Met Asn Phe Cys Cys Leu Gly Phe Ile Ala Tyr Ala Tyr  
 63 65 70 75 80  
 65 Ser Val Lys Ser Arg Asp Arg Lys Met Val Gly Asp Val Thr Gly Ala  
 66 85 90 95  
 68 Gln Ala Tyr Ala Ser Thr Ala Lys Cys Leu Asn Ile Ser Thr Leu Val  
 69 100 105 110  
 71 Leu Ser Ile Leu Met Val Val Ile Thr Ile Val Ser Val Ile Ile Ile  
 72 115 120 125  
 74 Val Leu Asn Ala Gln Asn Leu His Thr  
 75 130 135  
 77 <210> SEQ ID NO: 3  
 78 <211> LENGTH: 823  
 79 <212> TYPE: DNA  
 80 <213> ORGANISM: Mus musculus  
 82 <400> SEQUENCE: 3  
 83 ggatcacaga ctgactgcta attgggtctt ggttttaggt ctttcaaag actaaggcaat 60  
 84 cttgttccga gctagctttt gaggctctg cccatcgcat cgccatggag gaaccatcg 120  
 85 agaaaagtgcga cccaatgaag gaccctgaaa ctccctcagaa gaaagatgaa gaggacgctt 180  
 86 tggatgatac agacgtccta caaccagaaa cactagtaaa ggtcatgaaa aagctaacc 240  
 87 taaaaccccg tgcataagcg tccgcacgccc ggcgcagtct acggaaccgc attgcagccg 300  
 88 tacctgtgga gaacaagagt gaaaaaatcc ggagggaaatg tcaaagcgcc tttcccaaga 360  
 89 gaagggtccg cactttgttgc tcgggtctga aagaccctat agcaaagatg agaagacttg 420  
 90 ttccggattga gcagagacaa aaaaggctcg aaggaaatga gtttgaacgg gacagtgagc 480  
 91 cattcagatg tctctgcact ttctgcactt atcaaagatg ggatccctct gagaatgcga 540  
 92 aaatcggaa gaatttaggag cttacattgt acgctgcctt ggctgtcgac gatgccgcac 600  
 93 agcagatgtg aaagctattt ttgtttaag attaaacttt ttctggtgct gggaaatctt 660  
 94 aacttgtttaa cttttaattt gtagatagga tgacacaacga tccagattta tgtgaagttt 720  
 95 agaagcctca agctgtgagg cccaggcgtg aggaataaag taaatagaat ttggagtatg 780  
 96 tacgttctaa ttccagaaaa ttgtataaa aagcattttt gtt 823  
 98 <210> SEQ ID NO: 4  
 99 <211> LENGTH: 150  
 100 <212> TYPE: PRT  
 101 <213> ORGANISM: Mus musculus  
 103 <400> SEQUENCE: 4  
 104 Met Glu Glu Pro Ser Glu Lys Val Asp Pro Met Lys Asp Pro Glu Thr  
 105 1 5 10 15  
 107 Pro Gln Lys Lys Asp Glu Glu Asp Ala Leu Asp Asp Thr Asp Val Leu  
 108 20 25 30  
 110 Gln Pro Glu Thr Leu Val Lys Val Met Lys Lys Leu Thr Leu Asn Pro  
 111 35 40 45  
 113 Gly Val Lys Arg Ser Ala Arg Arg Arg Ser Leu Arg Asn Arg Ile Ala  
 114 50 55 60  
 116 Ala Val Pro Val Glu Asn Lys Ser Glu Lys Ile Arg Arg Glu Val Gln  
 117 65 70 75 80  
 119 Ser Ala Phe Pro Lys Arg Arg Val Arg Thr Leu Leu Ser Val Leu Lys  
 120 85 90 95  
 122 Asp Pro Ile Ala Lys Met Arg Arg Leu Val Arg Ile Glu Gln Arg Gln  
 123 100 105 110  
 125 Lys Arg Leu Glu Gly Asn Glu Phe Glu Arg Asp Ser Glu Pro Phe Arg

**RAW SEQUENCE LISTING**  
PATENT APPLICATION: US/10/646,390A

DATE: 11/15/2004  
TIME: 12:00:48

Input Set : A:\PTO.FG.txt  
Output Set: N:\CRF4\11152004\J646390A.raw

126	115	120	125
128	Cys Leu Cys Thr Phe Cys His Tyr Gln Arg Trp Asp Pro Ser Glu Asn		
129	130	135	140
131	Ala Lys Ile Gly Lys Asn		
132	145	150	
134	<210> SEQ ID NO: 5		
135	<211> LENGTH: 4925		
136	<212> TYPE: DNA		
137	<213> ORGANISM: Rattus sp		
139	<400> SEQUENCE: 5		
140	ccccccccc cccccccccc ccccccaccc cgacgtatga tggctcctag 60		
141	acgcacacacg aaggcgactc cccgcacatcat tcacgttagac ccgccttctg ctcccttgt 120		
142	cggggttttg ggaagcccg ggccctctc ttctcacctt gctccactag cacgcggctg 180		
143	ttttcaactga gcccagcact ggctaagtgg agcaccagga gtttcaggct atccttcaga 240		
144	ggccaagggtg tagtccatgg tgggctacag gagaccctct ctctccgtga gtacagagag 300		
145	gcaaacccaa gccagacagg ggtgtatgatt aggaacatac ctgcgtcggg gagaaaatac 360		
146	cggttcatat aggaataaga ggaaccagga ggtgttaag gctgtgggt ctgggtgcgg 420		
147	ggttttgac tctcaacaac cacgttcaga acgtgctgag ttttatgtat ggtgtagaat 480		
148	tcccttatca gcaattggc tccgcgggt ttctttttct ttttaattt ttaagtata 540		
149	atttgggtt tgaagcaact gtactggac tagaactccc tgtgtatcc agaatggaaat 600		
150	cccaaacttctt aggattaaag gttttagtgg gctgcagttt ggggtgggg ttgtttgtat 660		
151	tacgtttag cccaggctgg gctcaatctc aatccctctg cctctgcctt ctaaacgcta 720		
152	ggattaaaag tgctgcgcca tgatctgtt gtagtttat ttttattttt ttattttttt 780		
153	atttggctc ttttttttgg gagctgggaa ccgaaccggag ggccttgc ttccttaggca 840		
154	agcgctctac cactgagcta aatcccaac cccagtgttag ctttattttt aagaacagga 900		
155	gtcttgggtt tcaaaaacagt ttctctgttag ccctgggtt cctgaaactc cgtaaaccag 960		
156	gctgggttgg gactctgcctt taaaacact gggactaaag ggggttaccac ctccgtggc 1020		
157	tacaccggaa tcttttaagc ttcatgttcc cccgggtttt ttctttttt caccactt 1080		
158	ctggaaagcga tttctgttcc aaatttccat ttctggtaaa tgactctgag gggaaatagg 1140		
159	aacccagaat agattgagcc gggggctacc tgggaccctt cactccac ccccccagccg 1200		
160	ctgttgaagc tctttgcctt agggggctcc ggggttggata cctcttagca ctccgggtt 1260		
161	agggcgtggc tcgggaggag ccatttctt ggagaggaaa acaactgctg gccttgaatc 1320		
162	tgccttaata cctgacagtt acatgggacc ttcttatttc cacaggattt ttttagtctt 1380		
163	gtttgggaga ttttcaaatac ttgagactgc tcaacccttc ctggcctaact actcacaagg 1440		
164	ccaggctaga cccaaattct gtcaaccctt tctgtgttca aaacgggtggg tggctagctg 1500		
165	gctcacccctt ggtgttactt tgctttaaca ttctggaaaag ttgtggtaag ttctctgtat 1560		
166	aaaataggac catctactgg gtgtggccc atgtaaagca aggttggttt cccaaaatac 1620		
167	cctgtttaca tagatgtccg gaagcattgg agcaggtcaa tttagattttt gttggaaacag 1680		
168	cctgttttg gaaagcttc cagggcgaa aatgaacccca gaggcaactat tgggcaagcc 1740		
169	ctccggctaa gcaacacaat tggctgcagg ggtctctggaa agaggtgtga gacaagagag 1800		
170	aatatgcagg tttcaggacc tctgaaacttagt agttaggttgc ctgttacattt gtaacattgc 1860		
171	tgtaaagcaga acagccccatg gtaagaagct cagtggatct ctacaaacac taggatatct 1920		
172	gctcagggtt tatgaccagg ccctgttgcattt atgggttgc ttctgttgc ccctctttt 1980		
173	aagagggggtg attatctttt accccacttcc ttgtttctt ggggttattac ttgc当地 2040		
174	gcaaaaatgtt atacttactt aatgttccat tcttctgtt cagaaatctt acaaccagaa 2100		
175	acactagtaa aggtcatgaa aaagcttacc ctgttacccca gtgcaagcc gacaaaatat 2160		
176	catcgcttc aagggttcg tctccagggtt aagagccagc ctgtggagaa cagaagtgaa 2220		
177	agaatcatgaa gggaaatgttca aagccctt cccaggagaa gggccgcac tctgttgc 2280		
178	gtgctgaaag accccatagc aaggatgaga agatttgc gggtagttt cttgttgc 2340		

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/646,390A

DATE: 11/15/2004  
TIME: 12:00:48

Input Set : A:\PTO.FG.txt  
Output Set: N:\CRF4\11152004\J646390A.raw

179 cggggcatag atctaagagg aactctagcc tcaggaatgg cacctaggaa 2400  
 180 ttagacaag gatagtgact acctgtgatt cccagctcaa gaaaacaagc tccaaggcta 2460  
 181 tcctctactg cgcaagtctga agctgccag agctatatgc aaattgataa gtcagtataa 2520  
 182 catttatttt tgatatttca gactccctcc ccatagtc aactgccct ccagttcagt 2580  
 183 ccacggctct gcttctcccc cggtgctagg ctttgatgt ataaggctga cttagactgg 2640  
 184 atctcagagg tgaagtggac ctgttagtct ttgttagacca ggctggggtg gtttctgctt 2700  
 185 tctcagcgcc tagtcacat agtaggcatt ttaactttgt cttaatagta atttgatgaa 2760  
 186 ttttgtttt ctcttgaaga ttgagcagag acaaagacag cttgaaggaa atgaggtaaa 2820  
 187 tgcataatggg tgggttaggg gtctatggat gggtagggtg tcttggggactt 2880  
 188 tagacaaggg gtgtgtatgt ggagagttac cttctcaaca cagggaaatct ggttattaaa 2940  
 189 gcagtacttt aaaaataaaat aaaataaata aaataaaaaat aaagcagtag aaggggattt 3000  
 190 acatttcttt tgagttgcaa tatcctgatt aacatttttc tttcagagac gagatgagcc 3060  
 191 attcagatgt ctctgcactt tctgcattt tcagagatgg gatccttctg agaatgctaa 3120  
 192 aatcgggcag aaccagaaga attaggcag tttgaattgt acaccgtcct tgccgttaac 3180  
 193 ggtgccatgc agcagatgtg aaagctgttt ttttggtaa gattaaactt ttcttggc 3240  
 194 tggggaaatc tcttctaatt gctaaccctt aaattatata ggatgtgtga catttggatt 3300  
 195 catggaaatg acagatttac ccaagaattt agcatgagtc aaagcctggt agtttgattt 3360  
 196 agaaggtaat tggaaataat ctttttattt tagattttt agtttgcaga gaaatttgta 3420  
 197 ataaaaggcaa atttggttatc ttttataat acagaacaga tttagaatgag ccattggaga 3480  
 198 tggggactc gtttttaca ggtcatgtg tgggtgtgtg atgttcagag ttcaatgtgt 3540  
 199 gctaccctgt atttctgctt gaggcaaggt ctccatgagg cctagctggt ctaactccctg 3600  
 200 gtcctgcctt ttgtttccc ctgagttt acaccatagg cttgtcggca agatctggaa 3660  
 201 gaggcttgat gtttggatgtt gtgtgtgtt ataaaacaatt ggttgcata ttcctaaagt 3720  
 202 gtggcactgt attgacactgt ctgttcatg aggaagttaa tgaccggagc ataattgtat 3780  
 203 gctttatttc ctgagagaag tgtcaggaaa ggaggaggtt ggaagaaagc cccaggctgg 3840  
 204 ggttaagaggc actggctgtt tttccagagg tcctgagttc aattcccaggc aatcacctgg 3900  
 205 tggctcccgaa acatctgtta caggatccaa tggccctttt tgggtgtgtc aagaactccc 3960  
 206 taggcatttca gaggattttt gttttgttt tttttttttt tttttttttt ttcgtttttt 4020  
 207 tcagagctgg ggaaccgaac ccaggccctt gcgttgcata agcaagcgct ctaccactga 4080  
 208 gctaaatccc caaccctac aatggcttt ttctacctgc ttttgaatta tcaataaaag 4140  
 209 actggggcaa aagaaaggct ggagtgaatg agagagaaca tgtgaagagt aaatgagaga 4200  
 210 gagcatgagg gaatgaatga gagagtgaat gtgagaacga atgtgagagc gagtgagaga 4260  
 211 acatgagaag aacacqttaa gagtgagtga agagagaatg tgagggtgt atgaagattt 4320  
 212 tgtgtgggt tggggatttt gctcgttgtt agagtgttgc cctaggaagc acaaggccct 4380  
 213 gggtcggc cccagctcca aaaaaaagac caaaaaaaaa aaaaaaaaaa aaagattgtg 4440  
 214 tgtgtgtgtg aaaggagagt gcatgtgggt tgggtgttgc atgtcaagg tgggttatcaa 4500  
 215 gagtggtgtt gagagtgaaa ggtaatgaa cagagggtgtt catgagcgtt ggagtttgag 4560  
 216 aaaagaaaaac agcaataaaaa aaaaaaggcag agtgcacgag agaatgcaga gtgtgtgca 4620  
 217 cctcaagctg agacagagac agagagaagag agagagagag agagagactt taagccttga 4680  
 218 aattacctgt cagtttgcattt ccaaataatgtt gtgtgtgtt atttattttt agccttccag 4740  
 219 atccctgcctt ccagtggaga actctgatcc tatgttgagg ctggaccctg gcaatagtgg 4800  
 220 gcttcgttggaa aaatagtcaa aggaaacagt gctacaccat ggacttaagc cttagactc 4860  
 221 agttctggct tcaagagcag ctgtcagaaaa ataagtgtatg aactacttgc agtcgaactc 4920  
 222 gaatc 4925  
 224 <210> SEQ ID NO: 6  
 225 <211> LENGTH: 1444  
 226 <212> TYPE: DNA  
 227 <213> ORGANISM: Rattus sp  
 229 <400> SEQUENCE: 6

RAW SEQUENCE LISTING DATE: 11/15/2004  
 PATENT APPLICATION: US/10/646,390A TIME: 12:00:48

Input Set : A:\PTO.FG.txt  
 Output Set: N:\CRF4\11152004\J646390A.raw

230 ccaggattca gacgagctag gcctcatgca tggagacctt gcctcaagca gaaataaaaca 60  
 231 ggtagcaca cattgaactc tgaacatcac gagtgtgcac acacccacac atgcatctgt 120  
 232 aaaaaacgag tccccatctc caatggctcg ttctaatctg ttctgtgtat ttattaaaga 180  
 233 taacaaaattt gcctctatta caaatttetc tgcaaactag aaaatctaaa ataaaagatc 240  
 234 tattccaatt accttctaaa tcaaactacc gggctttgac tcatgctcaa ttcttggtt 300  
 235 aatctgtcat tccccatgaat ccaaatgtca cacatcttat ataatttaaa ggtagcaag 360  
 236 tagagatttc cccagcacca agaaaaagttt aatcttaaac aaaaaaaacag ctttcacatc 420  
 237 tgctgcattgg caccgttaac ggcaaggaca gtgtatgatt caaactgccc taattcttct 480  
 238 gggtctgccc aatttttagca ttctcagaag gatcccatt ctgataatgg cagaaagtac 540  
 239 agagacatct gaatggctca actcttctct catttccttc aagctgtctt tgtctctgtct 600  
 240 caatccgaac aaatcttctc atccttgcta tggggtctt cagcaccgac aacagtgtgc 660  
 241 ggacccctct cttgggaaag ggccttgaa cttccctcat gattcttca cttctgttct 720  
 242 ccacaggctg gctcttaatc tggagacgaa cccttgcac aagatgatat tttggccgat 780  
 243 tgagatagaa tatcaaaaaca acattnaaca tttaaataac ttaacgatat acacacctt 840  
 244 ttttttcca cttcccccaca cagacaaaaa acaaccctat ttttttta caaccccgcc 900  
 245 taagcaagcg aagcattagt aactgaccat tcatagaaag gaaacaccac cagaccacat 960  
 246 caaataaaat aaaatcaccc cccaaacccca cccctataaa aaacccgccc accacaccc 1020  
 247 atatactccc cccccccccc accatcaacta catcacccctc tccacccatt cccacacctt 1080  
 248 cccccaacat taaccccaacc ccacacggc aacccccaac accaacaat aaatttagaca 1140  
 249 catcgattt cataaattga cacaagaccc accccaaaag agcagcaag attagagcca 1200  
 250 catctcgcc ccaacacaat acactcaacc tgcatagttt ctatctccac cccaaacctt 1260  
 251 aaacaaaaat ctaatcagca ccaggcacc aagtatcactc cacactcaaa aacataccca 1320  
 252 ccaattaaac acgccccacc caccacaa cccacccgccc tgacaacaca cttcggaact 1380  
 253 accctcaaca tcacccaaag caatcgcaag ttacgatgac tccaaaccacc tcactctctc 1440  
 254 attg 1444  
 256 <210> SEQ ID NO: 7  
 257 <211> LENGTH: 7656  
 258 <212> TYPE: DNA  
 259 <213> ORGANISM: Rattus sp  
 261 <220> FEATURE:  
 262 <221> NAME/KEY: misc\_feature  
 263 <222> LOCATION: (7471)..(7471)  
 264 <223> OTHER INFORMATION: "n" is an unknown nucleotide  
 266 <220> FEATURE:  
 267 <221> NAME/KEY: misc\_feature  
 268 <222> LOCATION: (7554)..(7554)  
 269 <223> OTHER INFORMATION: "n" is an unknown nucleotide  
 271 <220> FEATURE:  
 272 <221> NAME/KEY: misc\_feature  
 273 <222> LOCATION: (7608)..(7608)  
 274 <223> OTHER INFORMATION: "n" is an unknown nucleotide  
 276 <400> SEQUENCE: 7  
 277 ctgcaagtag ttcatcattt acagatcaaa agaaaagaaga ataaaaaaaaac aaggtgtcat 60  
 278 gatccctcca aaagagtggc acacttcaac tgccagatcc aagatactga aatgggttagc 120  
 279 atgctggaga aagaatttca aagttagtta gagaatctgg ttgagcagag cacttgcatt 180  
 280 tcttccagag gatctgagtt caagttccag gaccttatatc acagtttct gtaactctag 240  
 281 ctccagaggg tctgacactt ctgttcaactg tgggcacctg cattcacacaga caaacataaa 300  
 282 gtatgttcatc accctttca cagaaaaccc acagcatgtg aggaaatccg ggtctctgcg 360  
 283 caatgccccca acagcagaag ggggagctg gagagatggt tcatctgtta gcccatttt 420

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/646,390A

DATE: 11/15/2004  
TIME: 12:00:49

Input Set : A:\PTO.FG.txt  
Output Set: N:\CRF4\11152004\J646390A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:7; N Pos. 7471,7554,7608  
Seq#:8; N Pos. 2115,2142,2143,2146

**VERIFICATION SUMMARY**  
PATENT APPLICATION: US/10/646,390A

DATE: 11/15/2004  
TIME: 12:00:49

Input Set : A:\PTO.FG.txt  
Output Set: N:\CRF4\11152004\J646390A.raw

L:401 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:7440  
M:341 Repeated in SeqNo=7  
L:467 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:2100